Radiative flux and forcing parameterization error in aerosol-free clear skies

Robert Pincus,[1] Eli J. Mlawer,[2] Lazaros Oreopoulos,[3] Andrew S. Ackerman,[4] Sunghye Baek,[5,6] Manfred Brath,[7] Stefan A. Buelhler,[7] Karen E. Cady-Pereira,[2] Jason N. S. Cole,[8] Jean-Louis Dufresne,[5] Maxwell Kelley,[4,9] Jiangnan Li,[8] James Manners,[10] David J. Paynter,[11] Romain Roehrig,[12] Miho Sekiguchi,[13] Daniel M. Schwarzkopf[11]

———

Corresponding author: Robert Pincus, 325 Broadway R/PSD1, Boulder CO 80305, USA. (Robert.Pincus@colorado.edu)

[1]University of Colorado and NOAA/Earth

This article reports on the accuracy in aerosol- and cloud-free conditions of the radiation parameterizations used in climate models. Accuracy is assessed relative to observationally-validated reference models for fluxes under present-day conditions and forcing (flux changes) from quadrupled concentrations of carbon dioxide. Agreement among reference models is typically within 1 W/m^2 , while parameterized calculations are roughly half as accurate in the longwave and even less accurate, and more variable, in the shortwave. Absorption of shortwave radiation is underestimated by most parameterizations in the present day and has relatively large errors in forcing. Error in present-day conditions is essentially unrelated to error in forcing calculations. Recent revisions to parameterizations have reduced error in most cases. A dependence on atmospheric conditions, including integrated water vapor, means that global estimates of parameterization error relevant for the radiative forcing of climate change will require much more ambitious calculations.

_____

System Research Lab, Physical Sciences

1. Assessing the accuracy of radiation parameterizations in climate models

Radiative transfer is unique among parameterization problems for global atmospheric models because the governing equations are deeply grounded in fundamental physics, the approximations (e.g. of one-dimensional radiative transfer) applicable across many relevant scales, and the result entirely deterministic. In aerosol-free clear skies, where scattering is small relative to absorption and emission, the problem is defined by the profile of extinction of the gaseous atmosphere, which is itself determined by the profiles of temperature, pressure, and the concentrations of radiatively-active gases. Fluxes of longwave or terrestrial radiation also depend on how the extinction profile is related to the profile of local emission (which also depends on temperature), while for fluxes of shortwave or solar radiation local emission can be neglected but Rayleigh scattering, which increases extinction and single-scattering albedo, changes the problem slightly. For monochromatic problems these calculations are straightforward so that the main challenge for atmospheric models is treating the spectral dependence of radiative fluxes.

The best available information for the spectral variation of radiation in the atmosphere comes from so-called line-by-line models with full spectral detail. When the state of the atmosphere is well-characterized such models are now able to match carefully-calibrated observations to within fractions of percent at full spectral resolution [see, for example, *Turner et al.*, 2004; *Alvarado et al.*, 2013]. This is sufficiently accurate, for example, that the very small spectrally-dependent signal from a decade's increase in CO_2 concentrations can be teased from surface measurements of spectral intensity that are dominated

Division, Boulder, Colorado, USA

by secular changes in temperature and water vapor [*Feldman et al.*, 2015]. The absolute

accuracy of these models in individual cases is difficult to judge because it is quite difficult

to separate errors in the characterization of the atmosphere from errors in the spectral

variation of optical thickness (vertically-integrated extinction). Estimates from remaining

spectrally-resolved differences against observations suggest that absolute errors in broad-

band fluxes (i.e. fluxes integrated over all energy emitted by the sun or earth) are no

larger than about 1 W/m^2 [*Oreopoulos and Mlawer*, 2010] in the longwave. Comparisons

in the shortave between spectrally-resolved measurements and model calculations have

been less extensive but suggest that broadband flux errors are somewhat larger.

Broadband fluxes are required by dynamical models of the atmosphere to compute

heating and cooling rates within the atmosphere and surface fluxes of energy. The opac-

ity of the cloud- and aerosol-free atmosphere [e.g. the 'clear, clean' conditions of *Ghan*,

2013] depends strongly enough on temperature and radiatively-active gas concentrations

that the extinction of the gaseous atmosphere is computed anew at each location and

time. Line-by-line models are impractically slow for this task because absorption coeffi-

cients due to gases vary by many orders of magnitude over very narrow spectral intervals.

Computationally-efficient parameterizations therefore use techniques such as exponen-

tial sum fitting of transmissivities [*Hunt and Grant*, 1969; *Wiscombe and Evans*, 1977],

correlated k-distributions [*Lacis and Oinas*, 1991; *Fu and Liou*, 1992], or the simplified ex-

change approximation [*Fels and Schwarzkopf*, 1981; *Schwarzkopf and Ramaswamy*, 1999]

----

[2]Atmospheric and Environmental

58 to approximate the spectral integral using a relatively small number of quadrature points

59 minimizing some measure of error across a distribution of atmospheric states. (Other

60 simplifications frequently used in parameterizations, such as low angular resolution or the

61 neglect of scattering in the longwave, don't introduce significant errors in aerosol-free clear

62 skies.)

63   Systematic assessments of parameterizations against the reference models on which they

64 are based date back more than two decades [*Ellingson and Fouquart*, 1991; *Cess et al.*,

65 1993]. Parameterizations have improved over time such that the best now reproduce

66 line-by-line flux calculations in present-day conditions to within about 1 W/m^2 [*Ore-*

67 *opoulos et al.*, 2012] although a surprisingly large range of error still exists for such a

68 well-understood problem. Parameterizations are substantially less accurate with respect

69 to forcing calculations [*Collins et al.*, 2006; *Forster et al.*, 2011] (i.e. the change in flux due

70 to a change in atmospheric composition) in part because the sets of atmospheric states

71 used to develop parameterizations often do not include large variations in composition

72 [see, e.g., section 4 of *Mlawer et al.*, 1997].

73   But changes in radiative flux that arise from increasing levels of greenhouse gases are

74 precisely the radiative forcing that drives climate change. Parameterization errors in

75 estimates of clear-sky instantaneous radiative forcing, a purely radiative quantity that

76 might be determined to high precision, are one reason that different models participating

77 in coordinated experiments [e.g. the Coupled Model Intercomparison Project or CMIP; see

78 *Taylor et al.*, 2012, describes the recently-concluded CMIP5] may be subject to different

_____

Research, Lexington, Massachusetts, USA

79 instantaneous forcing for the same change in atmospheric composition [*Zhang and Huang*,

80 2014; *Chung and Soden*, 2015]. This is especially true for large excursions from present-

81 day conditions such as "4×CO_2" (carbon dioxide concentrations quadrupled from their

82 pre-industrial value). Despite this concern abrupt 4×CO_2 experiments are widely used

83 to obtain estimates of sensitivity and effective radiative forcing from time-evolving flux

84 imbalance and temperature [*Gregory et al.*, 2004].

85 Here we examine the accuracy of the greenhouse gas component of a set of parameteri-

86 zations used in atmospheric models. We consider both fluxes under a range of present-day

87 conditions and the forcing from quadrupled concentrations of carbon dioxide under the

88 same range of conditions. We focus on parameterizations used in climate models (roughly,

89 those that participate in CMIP experiments) because these models are often used to un-

90 derstand the response to forcing; here we seek to identify errors in the forcing itself.

91 (Similar assessments by the chemistry-climate [*Forster et al.*, 2011] and aerosol [*Randles*

92 *et al.*, 2013] communities have focused on more process-specific metrics using a different

93 set of models, though with broadly similar results.) We demonstrate below that recent

94 revisions to parameterizations at several modeling centers have reduced error (at least in

95 our test cases), but that errors in absorption relative to surface and top-of-atmosphere

96 fluxes imply biases in present-day model hydrologic cycles and in the hydrologic sensitivity

97 to increasing greenhouse gas concentrations. We also show that parameterization accu-

98 racy often depends on atmospheric state and composition, suggesting that estimates of

————

[3]NASA Goddard Space Flight Center,

accuracy at the global scales needed to estimate radiative forcing error require large-scale assessments across a wide-ranging set of conditions.

2. Building on previous assessments

We report on calculations following a variant of the protocol used by the First Phase of the Continual Intercomparison of Radiation Codes [CIRC, see *Oreopoulos et al.*, 2012] in which broadband fluxes are computed for atmospheric and surface conditions (temperature, pressure, and greenhouse gas concentrations) specified in between 55 and 66 layers, roughly comparable to the vertical resolutions used in climate models. In most cases participants performed the calculations with radiative transfer codes extracted from the host model [see the recommendation by *Forster et al.*, 2011] although at least one implemented diagnostic calculations in a single-column version of the full model. We focus on the treatment of gases by considering four clear-sky cases, with precipitable water vapor ranging from 0.32 to 4.85 cm, originally chosen for good agreement between spectral measurements at the surface and reference model calculations using estimates of the overlying atmospheric state. We further simplify the problem by specifying no aerosols and spectrally-uniform surface albedo and emissivity. We perform two sets of experiments, one under the initially-observed atmospheric conditions [essentially repeating the shortwave 'case b' calculations of *Oreopoulos et al.*, 2012] and a second in which the observed concentrations of CO_2 have been quadrupled. Fluxes are reported at the top-of-atmosphere and the surface. Atmospheric absorption or emission (radiative divergence) is

---

Earth Science Division, Greenbelt,

118 computed as the difference between net-downward radiation at these two levels. Forcing

119 is the difference between $4\times CO_2$ and present-day conditions.

120   We solicited contributions from modeling centers that participate in CMIP exercises.

121 Many provided results from the current version of their radiation parameterization; some

122 also provided results from the version used in CMIP5 [*Taylor et al.*, 2012]. Calculations

123 made with version 12.2 of the LBLRTM model [*Clough et al.*, 2005] provide the benchmark,

124 as they did for CIRC (though CIRC used a slightly older version). Participating radiation

125 models and results from LBLRTM are detailed in the Supplementary Material. Although

126 our assessment does not include every parameterization used by models participating in

127 CMIP, several parameterizations are used in more than one model, so that the radiation

128 component is even less diverse than the already-interdependent set of models itself [*Knutti*

129 *et al.*, 2013].

3. Results from CMIP5 and beyond

130   As with past assessments, we find that reference models agree with one another to

131 within 1 W/m^2, or to fractions of a percent, for almost all calculations, though this

132 level of agreement is not a measure of uncertainty. The absolute accuracy of line-by-line

133 models, as measured by comparisons of high-resolution spectra against observations [e.g.

134 *Delamere et al.*, 2010; *Mlawer et al.*, 2012; *Alvarado et al.*, 2013], is improving over time. In

135 most cases improvements are due to refined information about the presence, position, and

136 strength of individual absorption lines or continua. Better spectral information introduces

137 changes in broadband fluxes, illustrated in Figure 1 by results from an older version (9.4+)

Maryland, USA

138 of LBLRTM (orange squares in this and all subsequent figures) which has less complete

139 spectroscopy and different models for the water vapor continuum, and agrees less well

140 with spectrally-resolved observations, than the more recent version (12.2) we use as our

141 benchmark. (Broadband fluxes and forcings are far more similar than spectrally-resolved

142 differences; see *Alvarado et al.* [2013].) More sophisticated treatment of line interactions is

143 also useful: the primary difference between the ARTS 2.2 (small black squares) and ARTS

144 2.3 (large black squares) is the treatment of line mixing by CO_2 (revisions to continua,

145 also present, have a small impact).

146   Nonetheless, line-by-line model results are almost entirely determined by the underlying

147 spectroscopic data including models for absorption continua. All the line-by-line models

148 in this study use the same continua in the longwave so spread among the reference models

149 does not reflect true uncertainty. Indeed, line-by-line models have agreed well with one

150 another for several decades [e.g. *Ridgway and Arking*, 1991], often to a greater degree

151 than they have agreed with observations.

152   Whatever the underlying uncertainty in line-by-line calculations, parameterizations are

153 constructed to reproduce these reference model results across a range of atmospheric con-

154 ditions. By this measure parameterization accuracies vary more widely across quantities

155 (e.g. top-of-atmosphere vs. surface flux), spectral intervals, and atmospheric conditions

156 than do reference models. For top-of-atmosphere longwave flux and forcing, as one exam-

157 ple, parameterization error is relatively low (Figure 1; see also *Oreopoulos et al.* [2012]),

—————

[4]Goddard Institute for Space Studies,

158   but even here errors of 2 W/m^2 (roughly 1%, depending on the case) in present-day net

159   flux and 0.5 W/m^2 (10%) in forcing from quadrupled CO_2 concentrations are common.

160   Figure 1 is meant to be illustrative; readers interested in details will find an exhaustive

161   set of figures in the Supplementary Material.

162   Revisions to parameterizations since CMIP5 have reduced errors, in some cases dra-

163   matically, with respect to both flux and forcing (compare the filled circles and dark grey

164   histogram bars in Fig. 1 to the open circles and light grey histogram bars; see also Fig-

165   ures S1 and S2). In most cases the reduced error is due to improved parameterizations

166   of spectroscopy i.e. more accurate treatment of how the optical thickness of the gaseous

167   component of the atmosphere depends on temperature, pressure, and composition. (Other

168   changes may be relevant for more complicated atmospheric conditions including aerosols

169   and clouds.)

170   The remaining error is not random, however. Many parameterizations systematically ei-

171   ther over- or under-estimate a given radiative quantity. Perhaps more importantly, errors

172   depend to some extent on atmospheric state including the profile of temperature and hu-

173   midity. This is evident in how the error in present-day flux in the GISS parameterizations

174   (shown in brown circles), and the error in forcing in all models but especially MSTRNX

175   (green circles), depend on column-integrated water vapor, which increases from left to

176   right in Fig. 1. As we discuss below, the state dependence of parameterization error

177   makes it hard to extrapolate from relatively few cases, even those constructed to sample,

New York, USA

178 say, latitudinal dependence [*Myhre and Stordal*, 1997; *Forster et al.*, 2011] to global or

179 regional estimates of error.

180 Relationships among errors at the top-of-atmosphere, surface, and within the atmo-

181 sphere can have important impacts on the host model: because globally-averaged surface

182 radiative fluxes affect the planet's surface temperature while net radiative cooling controls

183 the global amount of precipitation, radiation parameterization errors can impact present-

184 day energy and water cycles nearly independently. In general errors in longwave fluxes at

185 the top-of-atmosphere and surface often partially compensate for one another (Figure 2,

186 left panel). For some parameterizations the reasons for this compensation are understood:

187 the GFDL longwave parameterization, for example, includes energy in a narrower spectral

188 interval than do the reference calculations, resulting in smaller upwelling fluxes (positive

189 error) at both the surface and top of the atmosphere and hence very small impacts on

190 atmospheric cooling.

191 Calculations in the shortwave are generally less accurate in absolute terms than in

192 the longwave [see also *Oreopoulos et al.*, 2012], though for boundary fluxes the relative

193 errors are similar (within a few percent). More importantly, shortwave errors at the

194 top-of-atmosphere tend to be of opposite sign to those at the surface, leading to errors

195 in radiative divergence as large as 10 W/m^2 or 5% even for current parameterizations.

196 Errors for the GISS and LMDZ CMIP5-era parameterizations are even more dramatic. In

197 all but one parameterization shortwave absorption tends to be underestimated, implying

———

[5]CNRS/IPSL/LMD, Université Pierre et

a bias towards increased global precipitation in many host models. The exception is the

GFDL parameterization, for which slightly greater solar absorption is consistent with

the reference model on which the parameterization is based. This line-by-line model

Marie Curie, Paris, France

[6]Korea Institute of Atmosphere

Prediction Systems, Seoul, Korea

[7]Meteorological Institute, University of

Hamburg, Hamburg, Germany

[8]Canadian Center Climate Modelling &

Analysis, Environment Canada, Victoria,

British Columbia, Canada

[9]Trinnovim LLC, New York, New York,

USA

[10]Met Office, Exeter, UK

[11]NOAA Geophysical Fluid Dynamics

Laboratory, Princeton, New Jersey, USA

[12]Centre National de Recherches

Météorologiques-GAME, Météo-France &

CNRS, Toulouse, France

[13]Tokyo University of Marine Science and

Technology, Tokyo, Japan

absorbs more than our benchmark due to more recent underlying spectroscopic data and
to a different treatment of water vapor continua due to *Paynter and Ramaswamy* [2011].
Thus the overestimate of absorption by GFDL relative to our reference is more properly
viewed as a legitimate difference than an error.

Similarly, the partitioning of forcing errors within and at the boundaries of the atmo-
sphere means that parameterization errors may project differently onto climate (tempera-
ture) and hydrologic sensitivities to increasing concentrations of greenhouse gases. In the
shortwave most $4 \times CO_2$ forcing errors, like errors in present-day fluxes, are of opposite
sign at the boundaries (see the left panel of Figure 3) so that errors in radiative divergence
are relatively larger than at the surface or the top of the atmosphere. This implies that
radiation parameterization errors will have larger impacts on hydrologic sensitivity than
on climate sensitivity. The situation is much better in the longwave as shown in Figure S3.
In practice relationships between forcing at the boundaries of and within the atmosphere
will be modified by tropo- and stratospheric adjustments, normally considered as part of
the total forcing, that may dampen or amplify one error without affecting others.

It is tempting to think of parameterizations as being better or worse depending on
some measure of the error in, for example, present-day fluxes, but the size of the error
under present-day conditions says nearly nothing about the magnitude of errors in forcing
(Fig. 3, right panel, and Fig. S4). Instead, relationships between errors in present-
day and $4 \times CO_2$ conditions are determined more by the parameterization being used
than by atmospheric state. To the extent that these results hold on a global scale, it
suggests both that present-day biases do not imply errors in forcing (e.g. the Met Office

223 or GFDL parameterizations) and, conversely, that parameterizations with small errors

224 in the present day may still have significant forcing errors (e.g. the CCC and RRTMG

225 parameterizations).

4. Tracing error from a few aerosol-free clear skies to projections of climate change

226 A focus on clear skies may seem narrow given roughly two-thirds of the planet's surface

227 at any given time is covered by some kind of cloud [e.g. *Stubenrauch et al.*, 2013]. But clear-

228 sky errors exist even in cloudy skies (though they need not add linearly to errors introduced

229 in the treatment of clouds) and clouds do not affect forcing directly, so that characterizing

230 the error in aerosol-free clear skies remains a useful starting point for assessing errors in

231 radiative forcing.

232 The errors reported here come from snapshots at an essentially arbitrary point in time.

233 In particular, the models to be used in CMIP6 [*Meehl et al.*, 2014] are still under con-

234 struction and the radiation parameterizations are subject to revision before simulations

235 are made and shared. One goal of intercomparison exercises is to identify errors so they

236 may be fixed [see discussions in *Cahalan et al.*, 2005, for example] and this study has

237 already motivated two changes reflected in our results: spectroscopy was updated in the

238 GISS parameterizations and an important simplification (the assumption that surface

239 skin and surface air temperature are identical) was relaxed in the LMDZ and CNRM

240 longwave parameterizations after initial calculations demonstrated relatively large errors

241 in two cases. Further revisions may be expected in the next generation of climate models.

242 Some, we hope, will expand the range of atmospheres used to train the parameterization

243 to include conditions such as large concentrations of greenhouse gases.

244   The reference models on which radiation parameterizations are based are themselves

245 moving targets whose results change modestly as knowledge of the underlying spectroscopy

246 improves, especially in the shortwave. All reference models in this study, for example,

247 rely on the HITRAN spectroscopic data; differences between the 2012 and 2008 versions

248 [*Rothman et al.*, 2009, 2013] explains half or more (1 to 1.5 W/m^2, depending on the

249 case) of the difference in absorption calculations with LBLRTM (which uses the older

250 version) and the GFDL RFM line-by-line model. Periodic and/or continual assessments

251 of radiation parameterizations help modeling centers from straying too far from current

252 spectroscopic knowledge.

253   Radiation parameterization error is a function of the state of the atmosphere, as is

254 clear from Fig. 1. The set of conditions used in our calculations was originally de-

255 termined by the availability of high-quality observational data with which line-by-line

256 models could be validated [*Turner et al.*, 2004]. We cannot extrapolate from these

257 limited results to obtain global estimates of parameterization error in even in aerosol-

258 and cloud-free skies. One goal of the the Radiative Forcing Model Intercomparison

259 Project (RFMIP, `http://www.wcrp-climate.org/modelling-wgcm-mip-catalogue/`

260 `modelling-wgcm-mips/418-wgcm-rfmip`) associated with CMIP6 is to provide such

261 globally-relevant estimates, under present-day and future conditions and in strongly-forced

262 scenarios, by performing off-line radiative transfer calculations under a much larger range

263 of atmospheric conditions than is discussed here or in past assessments – a set of condi-

²⁶⁴ tions designed to allow for estimates of global parameterization error. These calculations

²⁶⁵ will enable the surprisingly large diversity in instantaneous clear-sky radiative forcing from

²⁶⁶ $4\times CO_2$ [*Zhang and Huang*, 2014; *Chung and Soden*, 2015], for example, to be decomposed

²⁶⁷ into correctable parameterization errors and legitimate differences in model climatology.

²⁶⁸   Highly accurate calculation of radiative fluxes is of primary interest to the climate

²⁶⁹ modeling community where long free-running simulations and highly variable greenhouse

²⁷⁰ gas concentrations are the norm. The range of parameterization errors, especially with

²⁷¹ respect to forcing, implies that specifying changes in greenhouse gas concentrations (as

²⁷² in CMIP) does not completely determine the instantaneous clear-sky radiative forcing to

²⁷³ which each model is subject. In fairness it is the effective radiative forcing, including

²⁷⁴ model-specific rapid adjustments [see, e.g. *Sherwood et al.*, 2015] and state-dependent

²⁷⁵ sensitivities, that is relevant to determining the long-term climate response. Diversity in

²⁷⁶ adjustments [*Andrews et al.*, 2012; *Zelinka et al.*, 2014] will partly mask and may well

²⁷⁷ outweigh the range of parameterization error. But rapid adjustments themselves can be

²⁷⁸ influenced by radiation parameterization error [*Ogura et al.*, 2014] and, given the maturity

²⁷⁹ of understanding about radiative transfer, global estimates of parameterization error are

²⁸⁰ a necessary first step in efforts to characterize and assess radiative forcing.

References

₂₉₃ Alvarado, M. J., V. H. Payne, E. J. Mlawer, G. Uymin, M. W. Shephard, K. E. Cady-

₂₉₄ Pereira, J. S. Delamere, and J. L. Moncet (2013), Performance of the Line-By-Line

₂₉₅ Radiative Transfer Model (LBLRTM) for temperature, water vapor, and trace gas re-

₂₉₆ trievals: Recent updates evaluated with IASI case studies, *Atmos. Chem. Phys.*, *13*(14),

₂₉₇ 6687–6711.

₂₉₈ Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor (2012), Forcing, feedbacks

₂₉₉ and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, *Geophys.*

₃₀₀ *Res. Lett.*, *39*(9).

₃₀₁ Cahalan, R. F., et al. (2005), The I3RC - Bringing together the most advanced radiative

₃₀₂ transfer tools for cloudy atmospheres, *Bull. Amer. Meteor. Soc.*, *86*, 1275+.

₃₀₃ Cess, R. D., et al. (1993), Uncertainties in Carbon Dioxide Radiative Forcing in Atmo-

₃₀₄ spheric General Circulation Models, *Science*, *262*(5137), 1252–1255.

₃₀₅ Chung, E.-S., and B. J. Soden (2015), An Assessment of Direct Radiative Forcing, Radia-

₃₀₆ tive Adjustments, and Radiative Feedbacks in Coupled Ocean–Atmosphere Models*, *J.*

Climate, *28*(10), 4152–4170.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown (2005), Atmospheric radiative transfer modeling: a summary of the AER codes, *J. Quant. Spectrosc. Radiat. Transfer*, *91*(2), 233–244.

Collins, W. D., et al. (2006), Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4), *J. Geophys. Res.*, *111*(D14), D14,317.

Delamere, J. S., S. A. Clough, V. H. Payne, E. J. Mlawer, D. D. Turner, and R. R. Gamache (2010), A far-infrared radiative closure study in the Arctic: Application to water vapor, *J. Geophys. Res.*, *115*(D17), D17,106.

Ellingson, R. G., and Y. Fouquart (1991), The Intercomparison of Radiation Codes in Climate Models: An Overview, *J. Geophys. Res.*, *96*(D5), 8925–8927.

Feldman, D. R., W. D. Collins, P. J. Gero, M. S. Torn, E. J. Mlawer, and T. R. Shippert (2015), Observational determination of surface radiative forcing by CO2 from 2000 to 2010, *Nature*, *519*(7543), 339–343.

Fels, S. B., and M. D. Schwarzkopf (1981), An efficient, accurate algorithm for calculating CO2 15 μm band cooling rates, *J. Geophys. Res.*, *86*(C2), 1205–1232.

Forster, P. M., et al. (2011), Evaluation of radiation scheme performance within chemistry climate models, *J. Geophys. Res.*, *116*(D10), D10,302.

Fu, Q., and K. N. Liou (1992), On the correlated *k*-distribution method for radiative transfer in nonhomogeneous atmospheres, *J. Atmos. Sci.*, *49*, 2139–2156.

329 Ghan, S. J. (2013), Technical Note: Estimating aerosol effects on cloud radiative forcing,

330   *Atmos. Chem. Phys.*, *13*(19), 9971–9974.

331 Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A.

332   Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative

333   forcing and climate sensitivity, *Geophys. Res. Lett.*, *31*(3), L03,205.

334 Hunt, G. E., and I. P. Grant (1969), Discrete Space Theory of Radiative Transfer and its

335   Application to Problems in Planetary Atmospheres, *J. Atmos. Sci.*, *26*(5), 963–972.

336 Knutti, R., D. Masson, and A. Gettelman (2013), Climate model genealogy: Generation

337   CMIP5 and how we got there, *Geophys. Res. Lett.*, *40*(6), 1194–1199.

338 Lacis, A. A., and V. Oinas (1991), A description of the correlated k-distribution method

339   for modeling non-grey gaseous absorption, thermal emission, and multiple scattering in

340   vertically inhomogeneous atmospheres, *J. Geophys. Res.*, *96*, 9027–9063.

341 Meehl, G. A., R. Moss, K. E. Taylor, V. Eyring, R. J. Stouffer, S. Bony, and B. Stevens

342   (2014), Climate Model Intercomparisons: Preparing for the Next Phase, *EOS Trans.*

343   *AGU*, *95*(9), 77–78.

344 Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), RRTM,

345   a validated correlated-k model for the longwave, *J. Geophys. Res.*, *103*, 16,663–16,682.

346 Mlawer, E. J., V. H. Payne, J. L. Moncet, J. S. Delamere, M. J. Alvarado, and D. C.

347   Tobin (2012), Development and recent evaluation of the MTCKD model of continuum

348   absorption, *Philosophical Transactions of the Royal Society A*, *370*(1968), 2520–2556.

349 Myhre, G., and F. Stordal (1997), Role of spatial and temporal variations in the compu-

350   tation of radiative forcing and GWP, *J. Geophys. Res.*, *102*(D10), 11,181–11,200.

Ogura, T., M. J. Webb, M. Watanabe, F. H. Lambert, Y. Tsushima, and M. Sekiguchi (2014), Importance of instantaneous radiative forcing for rapid tropospheric adjustment, *Climate Dyn.*, *43*(5-6), 1409–1421.

Oreopoulos, L., and E. Mlawer (2010), MODELING: The Continual Intercomparison of Radiation Codes (CIRC), *Bull. Amer. Meteor. Soc.*, *91*(3), 305–310.

Oreopoulos, L., et al. (2012), The Continual Intercomparison of Radiation Codes: Results from Phase I, *J. Geophys. Res.*, *117*(D6), D06,118.

Paynter, D. J., and V. Ramaswamy (2011), An assessment of recent water vapor continuum measurements upon longwave and shortwave radiative transfer, *J. Geophys. Res.*, *116*(D20), D20,302.

Randles, C. A., et al. (2013), Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: results from the AeroCom Radiative Transfer Experiment, *Atmos. Chem. Phys.*, *13*(5), 2347–2379.

Ridgway, W. L., and A. Arking (1991), Computation of atmospheric cooling rates by exact and approximate methods, *J. Geophys. Res.*, *96*(D5), 8969–8984.

Rothman, L. S., et al. (2009), The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, *110*(9–10), 533–572.

Rothman, L. S., et al. (2013), The HITRAN2012 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, *130*(0), 4–50.

Schwarzkopf, M. D., and V. Ramaswamy (1999), Radiative effects of CH4, N2O, halocarbons and the foreign-broadened H2O continuum: A GCM experiment, *J. Geophys. Res.*, *104*(D8), 9467–9488.

Sherwood, S. C., S. Bony, O. Boucher, C. Bretherton, P. M. Forster, J. M. Gregory, and B. Stevens (2015), Adjustments in the forcing-feedback framework for understanding climate change, *Bull. Amer. Meteor. Soc.*, *96*, 217–228.

Stubenrauch, C. J., et al. (2013), Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, *Bull. Amer. Meteor. Soc.*, *94*(7), 1031–1049.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An Overview of CMIP5 and the Experiment Design, *Bull. Amer. Meteor. Soc.*, *93*(4), 485–498.

Turner, D. D., et al. (2004), The QME AERI LBLRTM: A Closure Experiment for Downwelling High Spectral Resolution Infrared Radiance, *J. Atmos. Sci.*, *61*(22), 2657–2675.

Wiscombe, W. J., and J. W. Evans (1977), Exponential-sum fitting of radiative transmission functions, *J. Comp. Phys.*, *24*(4), 416–444.

Zelinka, M. D., T. Andrews, P. M. Forster, and K. E. Taylor (2014), Quantifying components of aerosol-cloud-radiation interactions in climate models, *J. Geophys. Res.*, *119*(12), 7599–7615.

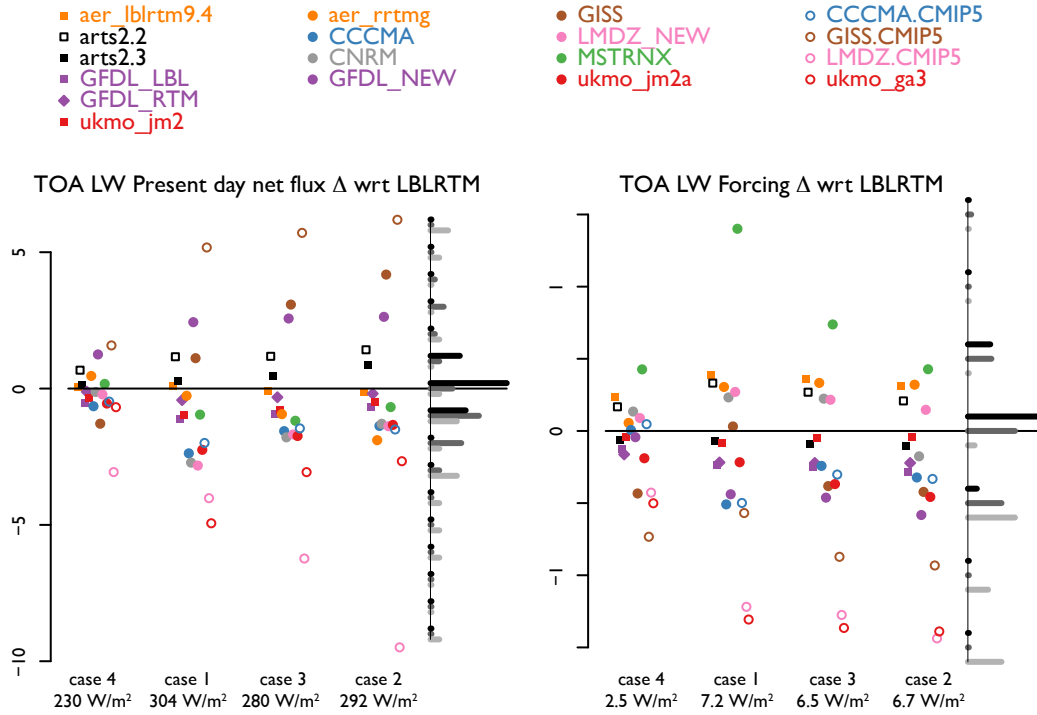Zhang, M., and Y. Huang (2014), Radiative Forcing of Quadrupling CO_2, *J. Climate*, *27*(7), 2496–2508.

Figure 1. Errors in top-of-atmosphere net downward broadband longwave flux under present-day conditions (left panel) and error in forcing from CO_2 concentrations quadrupled from present-day values (right), for a variety of reference models (squares, including one high-resolution k-distribution model) and parameterizations used in climate models (circles). Calculations in four atmospheric profiles are ordered by column-integrated water vapor ranging from 0.32 to 4.85 cm. Reference values are noted. Histograms along the right edge of each panel show the error by model type (black for reference models, dark grey for current parameterizations, light grey for parameterizations used in CMIP5; normalized across model categories) in intervals of 1 (left panel) or 0.5 (right) W/m^2) centered on 0. Reference models agree to within fractions of a percent in most calculations. Parameterization error is larger but still modest, with error in some parameterizations depending strongly on the water vapor path. Parameterizations that have been updated since CMIP5 are more accurate than their predecessors (compare filled circles and dark grey histogram bars to open circles and light grey bars.)
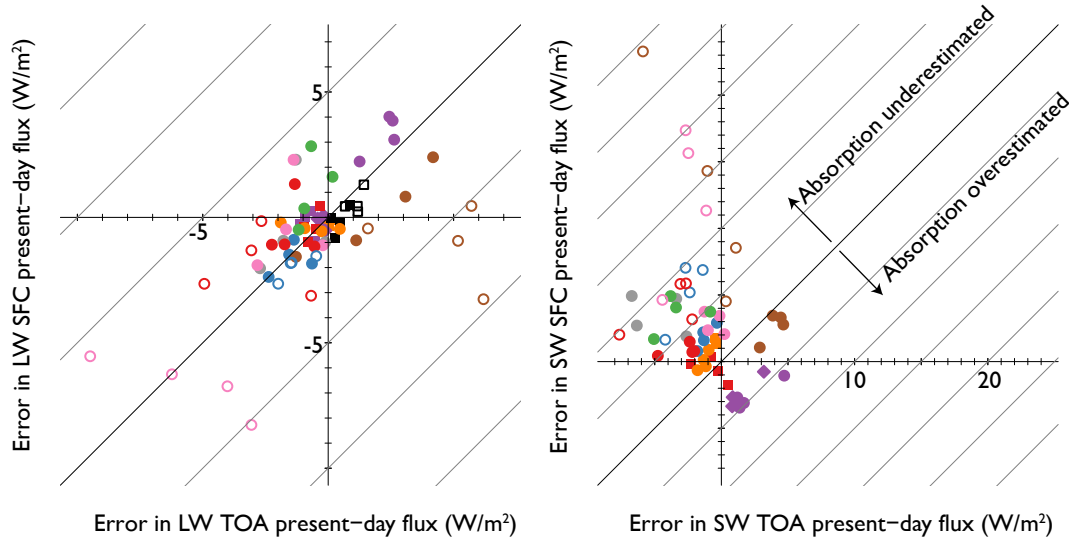
Figure 2.   Errors in surface net downward longwave (left panel) and shortwave (right) flux in the present day as a function of errors in net top-of-atmosphere flux, along with (diagonal) isolines of constant error in atmospheric absorption inferred from fluxes at the boundary. All four cases are shown. Axis tick marks are every 1 W/m² and absorption isolines every 5 W/m². Errors in longwave boundary fluxes and radiative divergence are generally within a few W/m², especially for current parameterizations, while errors in shortwave fluxes and absorption are twice as large.
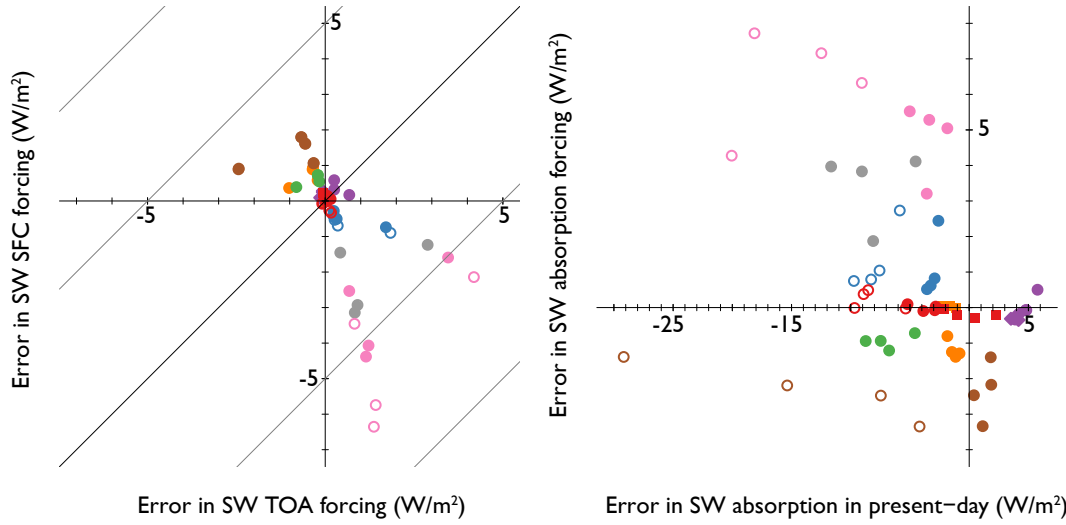
Figure 3. Left: Errors in surface net downward shortwave forcing from quadrupled present-day CO_2 concentrations as a function of error in net top-of-atmosphere forcing, along with (diagonal) isolines of constant error in atmospheric absorption forcing. Axis tick marks are every 1 W/m^2 and absorption isolines every 5 W/m^2. For many parameterizations errors in top-of-atmosphere forcing are smaller than errors in surface forcing, implying errors in absorption forcing with impacts on the host model's hydrologic cycle. Right: Error in solar absorption forcing as a function of error in the present-day estimates of solar absorption. Parameterizations have characteristic relationships among present-day and forcing errors that outweigh the impact of atmospheric state. Errors in the present-day are a poor predictor of errors in forcing.